ASYMPTOTICALLY OPTIMAL MAXIMUM-LIKELIHOOD ESTIMATION OF A CLASS OF CHAOTIC SIGNALS USING THE VITERBI ALGORITHM

David Luengo [1]*, Ignacio Santamaría* [2]*, Luis Vielva* [2]

[1] Dep. Teoría de la Señal y Comunicaciones (TSC), Universidad Carlos III de Madrid
Av. La Universidad 30, 28911, Leganés (Madrid), Spain
phone: + (34) 916248752, fax: + (34) 916249430, email: luengod@ieee.org
web: http://www.tsc.uc3m.es/

[2] Dep. Ingeniería de Comunicaciones (DICOM), Universidad de Cantabria
Av. Los Castros s/n, 39005, Santander, Spain
phone: + (34) 942201552, fax: + (34) 942201488, email: {nacho,luis}@gtas.dicom.unican.es
web: http://www.gtas.dicom.unican.es/

ABSTRACT

Chaotic signals and systems are potentially attractive in many signal processing and communications applications. Maximum likelihood (ML) and Bayesian estimators have been developed for piecewise-linear (PWL) maps, but their computational cost is excessive for practical applications. Several computationally efficient techniques have been proposed for this class of signals, but their performance is usually far from the optimum methods. In this paper, we present an asymptotically optimal estimator based on the Viterbi algorithm for estimating chaotic signals observed in additive white Gaussian noise. Computer simulations demonstrate that the performance of this estimator is similar to that of optimum methods with only a fraction of their computational cost.

1. INTRODUCTION

Chaotic signals (i.e. signals generated by a suitable nonlinear dynamical system in a chaotic state) have received much attention over the last decade. Although chaotic signals are generated by deterministic systems, they display features typical of purely random signals [1]: sensitivity to initial conditions, quickly decaying autocorrelation function, high bandwidth with an approximately flat spectral density, and practical unpredictability in the medium/long term.

These characteristics make them attractive in a wide range of signal processing and communications applications. In this paper we consider unidimensional chaotic maps. Although this is the simplest class of chaotic systems, they are useful in several different areas: random number generation [2], spread spectrum chaotic communications [3], watermarking [4], cryptography [5], etc. In any case, regardless of the application, it is necessary to develop computationally efficient detection and estimation algorithms which take into account the dual (deterministic/random) nature of chaotic signals and which show a robust behaviour under realistic conditions (e.g. in the presence of additive noise).

Estimation of chaotic signals has been addressed in several papers. The maximum likelihood (ML) estimator of the initial condition of a chaotic sequence has been developed in [6] for piecewise linear (PWL) maps. Bayesian estimators have been proposed as well for any PWL map in [7]. These estimators show a good performance and attain the Cramer-Rao lower bound (CRLB) asymptotically as the signal to noise ratio (SNR) goes to infinity.

Unfortunately, the computational cost of these optimum estimators grows exponentially with the length of the chaotic sequence, and cannot be reduced in general (in some particular cases, such as the tent-map, an efficient recursive implementation of the ML estimator is possible [8]). Consequently, there is still a need to obtain cost-effective estimators for their use in practical applications. Many suboptimal algorithms have been proposed (see for example [9, 10, 11, 12]), but their performance is usually far away from that of the ML and Bayesian estimators, specially in the low/medium SNR range.

In this paper we consider the use of the Viterbi algorithm, which in this case is an approximate method, for the estimation of the itinerary of the chaotic sequence. Once the itinerary is known, the ML estimator can be obtained in closed form for a PWL map, or through a simple local gradient descent or grid search method for non-PWL maps.

Note that the Viterbi algorithm has already been considered for the estimation of the itinerary of chaotic sequences in [13]. However, the work in [13] relies on a linear filter representation of the chaotic system which is not always possible, requires delay and truncation (thus generating pseudochaotic signals, which may lose some of the interesting features of the actual chaotic signals), and a trellis with a large number of states. Unlike [13], our approach is able to generate truly chaotic signals (since it is based on backward iteration of the chaotic system) is valid for any chaotic map, and provides a good performance with a reduced number of states.

2. CHAOTIC MAPS AND SYMBOLIC DYNAMICS

In this work we consider sequences generated by unidimensional chaotic maps. The n-th sample of the sequence is obtained iterating a known initial condition, $x[0]$, according to

$$x[n] = f(x[n-1]) = f^2(x[n-2]) = \ldots = f^n(x[0]), \quad (1)$$

where $f(x)$ is a suitable nonlinear and noninvertible function, $f^k(x)$ denotes the functional composition of $f(x)$ with itself k times, and $1 \leq n \leq N$. Although the estimation technique based on the Viterbi algorithm is valid for any chaotic map, in the sequel we concentrate on PWL maps defined on a phase space $I = [e_0, e_M]$, which can be described as

$$f(x) = \sum_{i=1}^{M} f_i(x) \; \mathbb{1}_{E_i}(x), \quad (2)$$

where $f_i(x) = a_i x + b_i$, $E_i = [e_{i-1}, e_i)$ for $0 \leq i \leq M-1$, $E_M = [e_{M-1}, e_M]$, and $_{E_i}(x)$ is an indicator or *characteristic function*, which denotes whether x belongs to a given region:

$$_{E_i}(x) = \begin{cases} 1, & x \in E_i; \\ 0, & x \notin E_i. \end{cases} \tag{3}$$

As an example, we focus on the *skew tent-map* (SK-TM),

$$f(x) = \begin{cases} \frac{x}{p}, & 0 \leq x < p; \\ \frac{1-x}{1-p}, & p \leq x \leq 1. \end{cases} \tag{4}$$

Where $0 < p < 1$ is a parameter of the map. In this case $M = 2$, $E_1 = [0,\ p)$ with $a_1 = 1/p$ and $b_1 = 0$, and $E_2 = [p,\ 1]$ with $a_2 = -1/(1-p)$ and $b_2 = 1/(1-p)$.

A very useful tool for analysing chaotic signals is *symbolic dynamics*. For any map we can define a partition of its phase space into a set of nonoverlapping intervals where the map is continuous and monotonous in such a way that they cover the whole phase space. This partition is never unique, but we can always find the simplest possible partition, which is called the natural or *generating partition* of the map. For PWL maps, this partition is clearly given by the E_i ($i = 1,\ \ldots,\ M$) and contains M elements. Thus, for the SK-TM it is simply $E_1 = [0,\ p)$ and $E_2 = [p,\ 1]$.

Now, we can define the *symbolic sequence* or *itinerary* of the map as the sequence of regions of its natural partition that the chaotic signal visits during its time evolution:

$$s[n] = i \Leftrightarrow x[n] \in E_i, \qquad n = 0,\ \ldots,\ N-1. \tag{5}$$

For PWL maps it can be easily shown that each point in their phase space has a unique symbolic sequence of length N associated. Moreover, a symbolic sequence of infinite length defines a single initial condition, $x[0]$, and an itinerary of finite length defines a closed region of possible initial values which becomes narrower as the length of the sequence increases [14].

Additionally, the symbolic sequence provides an alternative way of generating the chaotic signal. Instead of obtaining the n-th sample of the sequence, $x[n]$, iterating forward from a known initial condition, $x[0]$, we can obtain it iterating backwards from a known final condition, $x[N]$, as

$$x[n] = f^{-1}_{s[n]}(x[n+1]) = \ldots = f^{-(N-n)}_{s[n],\ \ldots,\ s[N-1]}(x[N]). \tag{6}$$

Where f^{-1} denotes the inverse map and $f^{-(N-k)}$ denotes the functional composition of f^{-1} with itself $N-k$ times. Note that (6) requires a priori knowledge of the itinerary in order to generate the desired chaotic sequence. For a generic PWL map the inverse function is

$$f^{-1}_{s[n]}(x) = \frac{x - b_{s[n]}}{a_{s[n]}}, \tag{7}$$

and for the SK-TM we have

$$f^{-1}_{s[n]}(x) = \begin{cases} px, & s[n] = 1; \\ 1 - (1-p)x, & s[n] = 2. \end{cases} \tag{8}$$

3. MAXIMUM LIKELIHOOD ESTIMATION

The data model that we consider in this paper is

$$\mathbf{y} = \mathbf{x} + \mathbf{w}, \tag{9}$$

where $\mathbf{y} = [y[0],\ \ldots,\ y[N]]^T$ is the observations vector, $\mathbf{x} = [x[0],\ \ldots,\ x[N]]^T$ is the chaotic sequence, and $\mathbf{w} = [w[0],\ \ldots,\ w[N]]^T$ is the noise vector whose samples, $w[n]$ ($0 \leq n \leq N$), correspond to zero-mean, white Gaussian noise with variance ^{2} (i.e. $\mathbf{w} \sim \mathcal{N}(\mathbf{0},\ ^2\mathbf{I})$, being $\mathbf{I}$ the $(N+1) \times (N+1)$ identity matrix).

The ML estimator obtains the chaotic sequence which maximizes the likelihood function, $p(\mathbf{y}; \mathbf{x})$, which is equivalent to minimizing the cost function,

$$J(\mathbf{x}) = (\mathbf{y} - \mathbf{x})^T (\mathbf{y} - \mathbf{x}), \tag{10}$$

since $p(\mathbf{y}; \mathbf{x})$ is a multivariate Gaussian PDF. However, it is apparent from (1) that the whole chaotic sequence can be expressed as a function of a single sample. Choosing $x[N]$ as a reference, we obtain an alternative cost function:

$$J(x[N], \mathbf{s}) = \sum_{k=0}^{N} \left(y[N-k] - f^{-k}_{\mathbf{s}_{N-k:N-1}}(x[N]) \right)^2, \tag{11}$$

where $\mathbf{s} = [s[0],\ \ldots,\ s[N-1]]^T$ is the symbolic vector (note that $s[N]$ is not used to generate the chaotic sequence), and $\mathbf{s}_{N-k:N-1} = [s[N-k],\ \ldots,\ s[N-1]]^T$. Then, the ML estimator of $x[N]$ is given by

$$\hat{x}_{\mathrm{ML}}[N] = \arg\min_{x[N]} J(x[N], \hat{\mathbf{s}}_{\mathrm{ML}}), \tag{12}$$

where $\hat{\mathbf{s}}_{\mathrm{ML}}$ is the ML estimate of $\mathbf{s}$.

We need to solve two problems to obtain the ML estimator of $x[N]$. First, the ML itinerary cannot be obtained taking derivatives of (11), because $J(x[N], \mathbf{s})$ is a discontinuous function of the symbolic sequence. Nevertheless, since the number of valid sequences is finite (at most M^N in general, 2^N for the SK-TM), a "brute force" approach is possible: test all the sequences, obtain their ML estimator, $\hat{x}^{\mathbf{s}_i}_{\mathrm{ML}}[N]$ ($i = 1,\ \ldots,\ M^N$), and select the one which minimizes (11).

A second problem comes from the fact that, even if we know the itinerary, we need a closed expression for the n-th iteration (backward in this case) of the chaotic sequence to be able to obtain $\hat{x}_{\mathrm{ML}}[N]$ in closed form. For PWL maps closed-form expressions have been developed in [6] and [15] for the forward and backward iteration respectively. These equations turn out to be linear in $x[0]$ and $x[N]$. Hence (11) is quadratic in $x[N]$ for a given itinerary, $\mathbf{s}_i$, and has a unique minimum, $\hat{x}_{\mathbf{s}_i}[N]$, which can be obtained easily taking its derivative with respect to $x[N]$ and equating it to zero [15].

However, $\hat{x}_{\mathbf{s}_i}[N]$ is the ML estimator of $x[N]$ only provided that the symbolic sequence is $\mathbf{s}_i$, and that $\hat{x}_{\mathbf{s}_i}[N]$ can generate a chaotic sequence with the specified itinerary (i.e. that $f^{-(N-k)}_{\mathbf{s}^i_{N-k:N-1}}(\hat{x}[N])$ exists and belongs to the phase space of the map, $I = [e_0, e_M]$, for $0 \leq k \leq N$). Otherwise it is necessary to apply a threshold to obtain the ML estimator of $x[N]$ for that symbolic sequence (see [6, 7] for a complete discussion in $x[0]$). Fortunately many chaotic maps, such as the SK-TM, are *onto* [14]: $f_i(x)$ maps E_i into the whole phase space

of the map for every i. Hence all the symbolic sequences are valid, and we only need to guarantee that $\hat{x}_{\mathbf{s}_i}[N]$ belongs to I:

$$\hat{x}_{\mathrm{ML}}^{\mathbf{s}_i}[N] = \begin{cases} e_0, & \hat{x}_{\mathbf{s}_i}[N] < e_0; \\ \hat{x}_{\mathbf{s}_i}[N], & \hat{x}_{\mathbf{s}_i}[N] \in I; \\ e_M, & \hat{x}_{\mathbf{s}_i}[N] > e_M. \end{cases} \tag{13}$$

Finally, the ML estimator of $x[N]$ is

$$\hat{x}_{\mathrm{ML}}[N] = \arg\min_i J(\hat{x}_{\mathrm{ML}}^{\mathbf{s}_i}, \mathbf{s}_i). \tag{14}$$

Once we obtain the ML estimator of the final condition of the sequence, $\hat{x}_{\mathrm{ML}}[N]$, and of the itinerary, $\hat{\mathbf{s}}_{\mathrm{ML}}$, the rest of the sequence can be obtained iterating backwards using (6):

$$\hat{x}_{\mathrm{ML}}[N-k] = f^{-k}_{\hat{\mathbf{s}}_{\mathrm{ML}}^{N-k:N-1}}(\hat{x}_{\mathrm{ML}}[N]), \qquad k = 1, \ldots, N. \tag{15}$$

4. ASYMPTOTIC MAXIMUM LIKELIHOOD ESTIMATION USING THE VITERBI ALGORITHM

The Viterbi decoding algorithm (VDA or VA) was developed originally by Andrew Viterbi in 1967 as an asymptotically optimum decoding method for convolutional codes, and was later extended by Omura and Forney to the detection of received signals distorted by intersymbol interference (ISI). First, a trellis diagram is constructed representing the valid transitions between states of the system at each iteration and their cost. Then, the VA searches for the shortest path through the trellis efficiently by merging paths and discarding unlikely sequences.

In this case, it is clear that a trellis diagram can be constructed for the chaotic sequence iterating backwards from $x[N]$, using the symbolic sequence, $\mathbf{s}$, to represent the states, and $s[n]$ for the transitions. However, it is also apparent that this trellis requires M^N states in general (2^N for the SK-TM) since the initial condition, $x[0]$, depends on the whole symbolic sequence (i.e. the system has a memory depth $N-1$).

In order to reduce the computational cost, we propose to use a trellis with a reduced set of states, $R = M^r$, and apply the VA. Although this is a suboptimal method, it provides a quasi-optimal performance, since far away symbols become less and less important in the estimation of $s[n]$ and $x[n]$.

The basic butterfly of the trellis diagram for $r = 1$ is shown in Fig. 1. The branch metrics are given by

$$c_{ij}[n] = |y[n+1] - f_j^{-1}(\hat{x}_i[n])|, \tag{16}$$

where $c_{ij}[n]$ is the cost of taking the j-th branch starting from the i-th node ($1 \le i, j \le R$) at the n-th time instant ($1 \le n \le N$), and $\hat{x}_i[n]$, $i \in \{1, \ldots, R\}$, is the sample obtained iterating backwards $N-n$ times from $x[N]$ using the best sequence which ends in the i-th node (state) at time n. The cost of the i-th node at the $(n+1)$-th instant can be obtained as usual from that of all the nodes at time n as

$$C_i[n+1] = \min_{1 \le j \le R} \{C_j[n] + c_{ji}[n]\}. \tag{17}$$

Since $x[N]$ is not known, as the starting sample of each state we use the closest point in E_i to $y[N]$:

$$\hat{x}_i[N] = \begin{cases} e_{i-1}, & y[N] < e_{i-1}; \\ y[N], & y[N] \in E_i; \\ e_i, & y[N] > e_i. \end{cases} \tag{18}$$

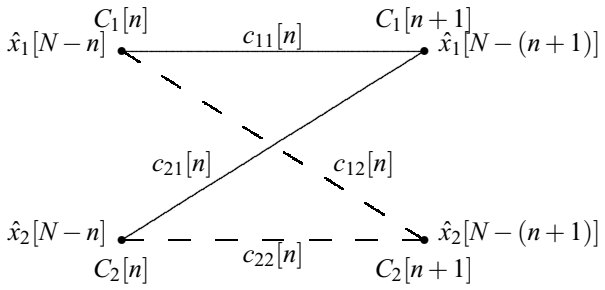

Figure 1: Basic butterfly of the trellis for the SK-TM using only two states per iteration ($M = 2$, $r = 1$).

Finally, $x[N]$ is estimated using the ML estimator given by (14) with the itinerary obtained with the VA, and the rest of the sequence is generated applying (15).

5. SIMULATION RESULTS

In this section we analyze the performance of the ML estimator based on the Viterbi algorithm. We first study short sequences, namely with $N = 4$. For the first example we consider an SK-TM with parameter $p = 0.2$, and an initial condition $x[0] = 0.1934$. Fig. 2 shows the mean square error (MSE) of the sequence obtained averaging 1000 simulations for all the estimators considered: the exact ML estimator, the VA based estimator, and the hard-censoring ML (HC-ML) estimator, which constructs an itinerary by hard-censoring of the noisy data and then applies the ML estimator [12]. The VA estimator attains the CRLB at the same SNR than the exact ML estimator, and provides a similar performance. Both the ML and the VA estimators provide a highly superior performance than the HC-ML estimator.

The exact SNR at which the CRLB is attained depends greatly on the chaotic map and the initial condition. As a second example we consider the alternative map given by

$$f(x) = \begin{cases} \frac{x}{p}, & 0 \le x < p; \\ \frac{x-p}{1-p}, & p \le x < 1; \end{cases} \tag{19}$$

for which $0 < p < 1$. For $p = 0.5$ this map is usually known as the Bernouilli or *binary shift map* (BSM), and is much

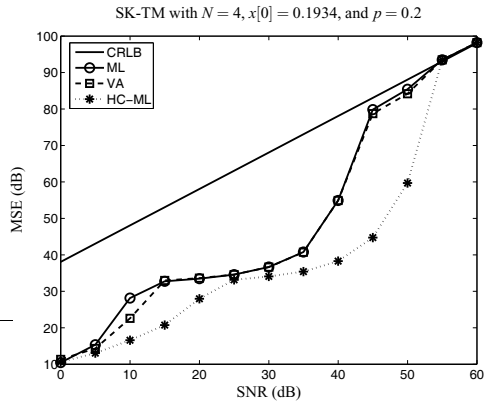

Figure 2: MSE for the SK-TM of the considered estimators: ML, VA, and HC-ML.

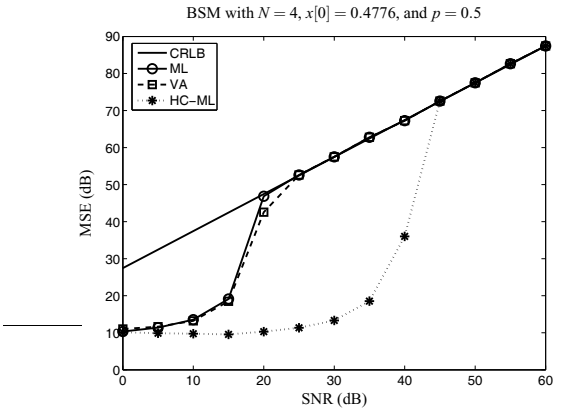

Figure 3: MSE for the BSM of the considered estimators: ML, VA, and HC-ML.

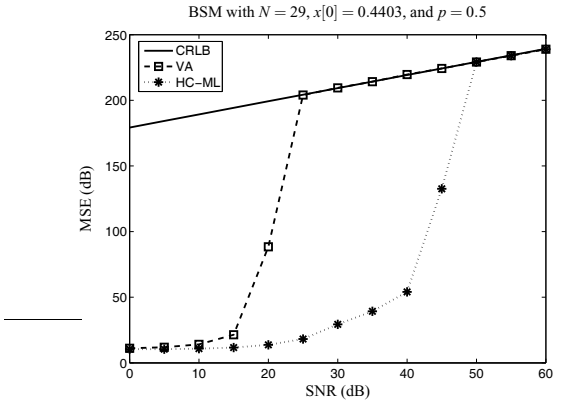

Figure 4: MSE for the BSM of the considered estimators: VA and HC-ML.

easier to estimate than the SK-TM, as Fig. 3 demonstrates. Now the ML estimator attains the CRLB at a greatly reduced SNR (25 dB compared to the 55 dB of Fig. 3). The VA achieves a quasi-optimal performance again and provides a huge improvement with respect to the HC-ML.

Finally, we consider the estimation of longer sequences. Specifically, we have used the BSM with $N = 29$ and $p = 0.5$. In this case the ML becomes unfeasible to calculate since it would require exploring 2^{29} (over $500 \cdot 10^6$) itineraries and thus the VA and HC-ML estimators are compared in Fig. 4. Once more the VA improves greatly the performance of the HC-ML estimator (e.g. it attains the CRLB at 20 dB of SNR compared to the 50 dB required by the HC-ML).

6. CONCLUSIONS

The ML estimator of chaotic signals exhibits an exponential increase in its computational cost with the length of the sequence. In this paper we have developed an efficient estimator based on the Viterbi algorithm which achieves a quasi-optimal performance with a reduced computational cost. Future lines of research include extending the method to non-PWL maps and maximum a posteriori (MAP) estimation, and achieving an additional complexity reduction combining the VA with a sphere decoding algorithm as in [16].

Acknowledgement

This work has been partially supported by MCYT (Ministerio de Ciencia y Tecnología) under grant TEC2004-06451-C05-02.

REFERENCES

[1] K. S. Chan and H. Tong, *Chaos: A Statistical Perspective*. New York: Springer-Verlag, 2001.

[2] S. Callegari, R. Rovatti, and G. Setti, "Embeddable ADC-based true random number generator for cryptographic applications exploiting nonlinear signal processing and chaos," *IEEE Trans. on Signal Processing*, vol. 53, no. 2, pp. 793–805, Feb. 2005.

[3] F. C. M. Lau and C. K. Tse, *Chaos-Based Digital Communication Systems*. Berlin: Springer-Verlag, 2003.

[4] A. Tefas, A. Nikolaidis, N. Nikolaidis, V. Solachidis, S. Tsekeridou, and I. Pitas, "Performance analysis of correlation-based watermarking schemes employing Markov chaotic sequences," *IEEE Trans. on Signal Processing*, vol. 51, no. 7, pp. 1979–1994, Jul. 2003.

[5] Y. Hwang and H. C. Papadopoulos, "Physical-layer secrecy in AWGN via a class of chaotic DS/SS systems: Analysis and design," *IEEE Trans. on Signal Processing*, vol. 52, no. 9, pp. 2637–2649, Sep. 2004.

[6] C. Pantaleón, D. Luengo, and I. Santamaría, "Optimal estimation of chaotic signals generated by piecewise-linear maps," *IEEE Signal Processing Letters*, vol. 7, no. 8, pp. 235–237, Aug. 2000.

[7] C. Pantaleón, L. Vielva, D. Luengo, and I. Santamaría, "Bayesian estimation of chaotic signals generated by piecewise-linear maps," *Signal Processing*, vol. 83, pp. 659–664, Mar. 2003.

[8] H. C. Papadopoulos and G. W. Wornell, "Maximum-likelihood estimation of a class of chaotic signals," *IEEE Trans. on Information Theory*, vol. 41, no. 1, pp. 312–317, Jan. 1995.

[9] S. M. Kay and V. Nagesha, "Methods for chaotic signal estimation," *IEEE Trans. on Signal Processing*, vol. 43, no. 8, pp. 2013–2016, Aug. 1995.

[10] L. Cong, W. Xiaofu, and S. Songgeng, "A general efficient method for chaotic signal estimation," *IEEE Trans. on Signal Processing*, vol. 47, no. 5, pp. 1424–1428, May 1999.

[11] S. Wang, P. C. Yip, and H. Leung, "Estimating initial conditions of noisy chaotic signals generated by piece-wise linear Markov maps using itineraries," *IEEE Trans. on Signal Processing*, vol. 47, no. 12, pp. 3289–3302, Dec. 1999.

[12] C. Pantaleón, D. Luengo, and I. Santamaría, "Optimal estimation of a class of chaotic signals," in *Proc. 16th World Computer and Communications Conf. - Int. Conf. on Signal Processing (WCCC-ICSP)*, vol. 1, 2000, pp. 276–280.

[13] M. Ciftci and D. B. Williams, "Optimal estimation and sequential channel equalization algorithms for chaotic communications systems," *EURASIP Journal on Applied Signal Processing*, vol. 4, pp. 249–256, 2001.

[14] R. L. Devaney, *An Introduction to Chaotic Dynamical Systems*. Reading, MA (USA): Perseus Books, 1989.

[15] D. Luengo, C. Pantaleón, and I. Santamaría, "Competitive chaotic AR(1) model estimation," in *Proc. IEEE Neural Networks for Signal Processing (NNSP) Work.*, 2001, pp. 83–92.

[16] H. Vikalo, B. Hassibi, and U. Mitra, "Sphere-constrained ML detection for frequency-selective channels," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, vol. IV, Hong Kong (China), Apr. 6–10 2003, pp. 1–4.